EVALUATION PRINCIPLES AND PRACTICES

AN INTERNAL WORKING PAPER

THE WILLIAM AND FLORA HEWLETT FOUNDATION

Prepared by:
Fay Twersky
Karen Lindblom

December 2012

TABLE OF CONTENTS

Cover image: *Measuring Infinity* by Jose de Rivera at the Smithsonian Museum of American History

EVALUATION IS PART OF THE FABRIC OF THE WILLIAM AND FLORA HEWLETT Foundation. It is referenced in our guiding principles. It is an explicit element of our outcome-focused grantmaking. And evaluation is practiced with increasing frequency, intensity, and skill across all programs and several administrative departments in the Foundation.

The purpose of this document is to advance the Foundation's existing work so that our evaluation practices become more consistent across the organization. We hope to create more common understanding of our philosophy, purpose, and expectations regarding evaluation as well as clarify staff roles and available support. With more consistency and shared understanding, we expect less wheel re-creation across program areas, greater learning from each other's efforts, and faster progress in designing meaningful evaluations and applying the results.

The following paper is organized into four substantive sections: (1) Principles, (2) Organizational Roles, (3) Practice Guide, and (4) Special Evaluation Cases. Supporting documents include a glossary of terms (Appendix A). The Principles and Organizational Roles should be fairly enduring, while the Practice Guide should be regularly updated with new examples, tools, and refined guidance based on lessons we learn as we design, implement, and use evaluations in our work.[1]

> **Hewlett Foundation Guiding Principle #3:**
>
> *The Foundation strives to maximize the effectiveness of its support.*
>
> This includes the application of outcome-focused grantmaking and the practice of evaluating the effectiveness of our strategies and grants.

> **What Is Evaluation?**
>
> Evaluation is an independent, systematic investigation into how, why, and to what extent objectives or goals are achieved. It can help the Foundation answer key questions about grants, clusters of grants, components, initiatives, or strategy.
>
> **What Is Monitoring?**
>
> Grant or portfolio monitoring is a process of tracking milestones and progress against expectations, for purposes of compliance and adjustment. Evaluation will often draw on grant monitoring data but will typically include other methods and data sources to answer more strategic questions.

---

[1] *While we appreciate the interconnectedness of strategy, monitoring, organizational effectiveness, and evaluation, this paper does NOT focus on those first three areas.* Those processes have been reasonably well defined in the Foundation and are referenced, as appropriate, in the context of evaluation planning, implementation, and use.

History

Recently, the Foundation adopted a common strategic framework to be used across all its program areas: Outcome-focused Grantmaking (OFG).[2] Monitoring and evaluation is the framework's ninth element, but expectations about what it would comprise have not yet been fully elaborated. Some program teams have incorporated evaluation at the start of their planning, while others have launched their strategies without a clear, compelling evaluation plan.

The good news is that, two to three years into strategy implementation, these programs typically have commissioned generally useful evaluations. The bad news is that they likely missed important learning opportunities by start-ing evaluation planning late in the process. Bringing evaluative thinking and discipline to the table early and often helps sharpen a strategy by clarifying assumptions and testing the logic in a theory of change. Early evaluation plan-ning also helps avoid the penalties of a late start: (1) missing a "baseline"; (2) not having data available or collected in a useful common format; (3) surprised, unhappy, or unnecessarily burdened grantees; and (4) an initiative not opti-mally designed to generate the hoped-for knowledge.

Based on these lessons of recent history, we are adapting our evaluation prac-tice to optimize learning within and across our teams. Staff members are eager for more guidance, support, and opportunities to learn from one another. They are curious, open-minded, and motivated to improve. Those are terrific attri-butes for an evaluation journey, and the Foundation is poised to productively focus on evaluation at this time.

This paper is the result of a collaborative effort, with active participation from a cross-Foundation Evaluation Working Group. Led by Fay Twersky and Karen Lindblom, members have included Paul Brest, Susan Bell, Barbara Chow, Ruth Levine, John McGuirk, Tom Steinbach, Jen Ratay, and Jacob Harold.

Intended Audience

Originally, this paper's intended audience was the Hewlett Foundation's staff—present and future. And of course, the *process of preparing the paper*, of involving teams and staff across the Foundation in fruitful conversation and skill building, has been invaluable in perpetuating a culture of inquiry and practical evalu-ation. Since good evaluation planning is not done in a vacuum, we asked a sample of grantees and colleagues from other foundations to offer input on an earlier draft. They all encouraged us to share this paper with the field, as they found it to be "digestible" and relevant to their own efforts.

While our primary audience remains Foundation staff, we now share the paper broadly, not as a blueprint, but in a spirit of collegiality and an interest in con-tributing to others' efforts and continuing our collective dialogue about evalua-tion practice.

---

[2] See the Hewlett Foundation's OFG memo for a complete description of this approach.

We aspire to have the following principles guide our evaluation practice:

1. **We lead with purpose. We design evaluation with actions and decisions in mind. We ask, "How and when will we use the information that comes from this evaluation?"** By anticipating our information needs, we are more likely to design and commission evaluations that will be useful and used. It is all too common in the sector for evaluations to be commissioned without a clear purpose, and then to be shelved without generating useful insights. We do not want to fall into that trap.

2. **Evaluation is fundamentally a learning process. As we engage in evaluation planning, implementation, and use of results, we actively learn and adapt.** Evaluative thinking and planning inform strategy development and target setting. They help clarify evidence and assumptions that undergird our approach. As we implement our strategies, we use evaluation as a key vehicle for learning, bringing new insights to our work and the work of others.

3. **We treat evaluation as an explicit and key part of strategy development.** Building evaluative thinking into our strategy development process does two things: (1) it helps articulate the key assumptions and logical (or illogical) connections in a theory of change; and (2) it establishes a starting point for evaluation questions and a proposal for answering them in a practical, meaningful sequence, with actions and decisions in mind.

7 Principles of Evaluation

1. We lead with purpose.

2. Evaluation is a learning process.

3. Evaluation is an explicit and key part of strategy development.

4. We strategically choose *what* to evaluate.

5. We choose methods that maximize rigor without compromising relevance.

6. We share our findings with appropriate audiences.

7. We use the data!

4. **We cannot evaluate everything, so we choose strategically.** Several criteria guide decisions about where to put our evaluation dollars, including the opportunity for learning; any urgency to make course corrections or future funding decisions; the potential for strategic or reputational risk; size of investment as a proxy for importance; and the expectation of a positive expected return from the dollars invested in an evaluation.

5. **We choose methods of measurement that allow us to maximize rigor without compromising relevance.** We seek to match methods to questions and do not routinely choose one approach or privilege one method over others. We seek to use multiple methods and data sources when possible in order to strengthen our evaluation design and reduce bias. All evaluations clearly articulate methods used and their limitations.
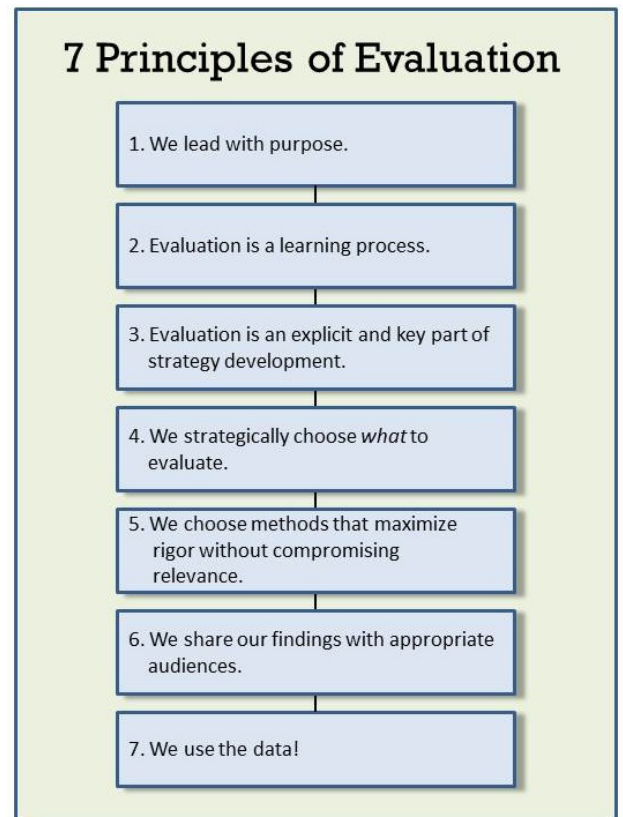
6. **We share our intentions to evaluate, and our findings, with appropriate audiences.** As we plan evaluations, we consider and identify audiences for the findings. We communicate early with our grantees and co-funders about our intention to evaluate and involve them as appropriate in issues of design and interpretation. We presumptively share the results of our evaluations so that others may learn from our successes and failures. We will make principled exceptions on a case-by-case basis, with care given to issues of confidentiality and support for an organization's improvement.

7. **We use the data!** We take time to reflect on the results, generate implications for policy or practice, and adapt as appropriate. We recognize the value in combining the insights from evaluation results with the wisdom from our own experiences. We support our grantees to do the same.

{ *We seek to maximize rigor without compromising relevance.* }

As the Foundation develops more formal systems and guidance for our evaluation work, it is appropriate to clarify basic expectations and roles for staff. As this work matures, and as our new central evaluation function evolves, we will continue to identify the best approaches to evaluation and refine these expectations accordingly.

Although we address the amount of time and effort staff may be expected to give to this work, it is important to note that the Foundation is less interested in the number of evaluations than in their high quality. Our standards are defined in the principles above and also informed by our practical learning and application of lessons.

Program and Operational Staff

Program and relevant operational staff (e.g., in the Communications and IT departments) are responsible and accountable for designing, commissioning, and managing evaluations, as well as for using their results. Programs are free to organize themselves however they deem most effective to meet standards of quality, relevance, and use. They may use a fully distributed model, with program officers responsible for their own evaluations, or they may designate a team member to lead evaluation efforts.

{ *Program and operational staff have primary responsibility for the evaluations they commission.* }

At least one staff member from each program will participate in a cross-Foundation Evaluation Community of Practice in order to support mutual learning and build shared understanding and skills across the organization. This participant could be a rotating member or standing member.

As part of programs' annual Budget Memo process and mid-course reviews, staff will summarize and draw on both monitoring and evaluation data—providing evidence of what has and has not worked well in a strategy and why. Staff are expected to use this data analysis to adapt or correct their strategy's course.

In general, program officers will spend 5 to 20 percent of their time designing and managing evaluations and determining how to use the results. This overall expectation is amortized over the course of each year, though of course there are periods when the time demands will be more or less intensive.

- The most intensive time demands tend to occur at the beginning and end of an evaluation—that is, when staff are planning and then using results.
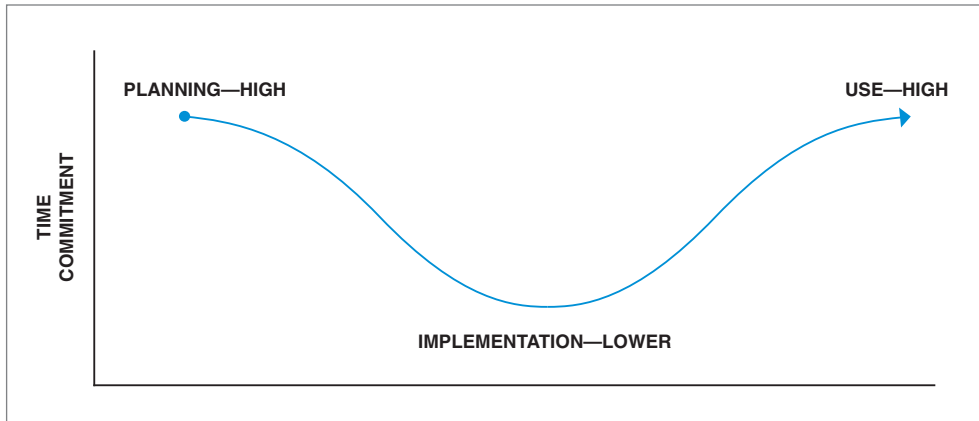
During these periods, full days can be devoted to the evaluation. For instance, planning requires considerable time to clarify design, refine questions, specify methods, choose consultants, and set up contracts. During use, staff spend time meeting with consultants, interpreting results, reviewing report drafts, communicating good or bad news, and identifying implications for practice.

- Less staff time is usually required during implementation, while evaluators are collecting data in the field. Ongoing management of their work takes some time, but, on the whole, not as much.

In general, program officers are expected to effectively manage one significant evaluation at any given time (maybe two, under the right circumstances). This includes proper oversight at each stage, from design through use and sharing of the results. When planning how to share results broadly, program staff should consult with the Foundation's Communications staff about the best approach.

Central Evaluation Support

As our approach to evaluation has become more deliberate and systematic, the Foundation's leadership has come to appreciate the value and timeliness of expert support for this work across the organization. Therefore, as part of its new Effective Philanthropy Group, the Foundation is creating a central support function for programs' evaluation efforts. It will:

- Provide consultation during strategy development, including teasing out assumptions and logical underpinnings in the theory of change.

- Support program staff in framing evaluation priorities, questions, sequencing, and methods. Help develop Requests for Proposals (RFPs) and review proposals.

- Maintain updated, practical, central resources: a vetted list of consultants with desired core competencies; criteria for assessing evaluation proposals; and examples of evaluation planning tools, RFPs, and

{ Central evaluation support is oriented toward consultation, NOT compliance. }

evaluation reports, including interim reports, internal and external reports, and executive summaries. Coordinate with the Foundation's Organizational Learning staff.

- Develop, test, and support the implementation of an application template and workflow for evaluation grants, including grant agreement letters. Coordinate with the relevant Foundation administrative departments: Grants Management and Legal.

- Provide or broker evaluation training for program staff in different formats (e.g., internal workshops, on-the-job training and coaching, and referrals to external resources, as appropriate).

- Spearhead an internal Evaluation Community of Practice for program staff who are leading evaluation efforts in their teams and want to share and deepen their skills and knowledge.

- Support external sharing of results as appropriate—coordinating with relevant program, Legal and Communications staff as well as grantees and other external partners.

- Work with Human Resources to refine job descriptions and performance review tools to accurately reflect evaluation responsibilities.

- Debrief every evaluation with the appropriate program staff: what went well, what didn't, key lessons, and actions taken as a result. Synthesize and share relevant lessons with other program staff so they can benefit from promising practice and lessons learned.

- Position the Foundation as a leader in the philanthropic evaluation field, in close coordination with Communications staff. Stay current with and contribute to the state of the art of evaluation.

- Coordinate as needed with the Human Resources, Organizational Learning, Philanthropy Grantmaking, and Organizational Effectiveness staff on any overlapping areas of learning, assessment, and training—both for Foundation staff and grantees.

Organizational Checks and Balances

How do we ensure that the Foundation does not simply commission evaluations that give us the answers we want? The practice guide that follows outlines a number of steps we are taking including: (1) building evaluation in from the beginning of a strategic initiative; (2) involving our board of directors in articulating key evaluation questions and then circling back with answers when we have them; (3) requiring methodology be clearly articulated for every evaluation—methodology that maximizes both rigor and relevance; (4) providing central expertise to review evaluation designs, proposals, and help interpret findings; (5) considering alternative explanations when interpreting results; and (6) debriefing every evaluation experience with a central evaluation officer—on all relevant lessons—to guard against easy answers or ignoring key findings.

This Practice Guide follows the three stages of evaluation: (1) planning, (2) implementation, and (3) practical use of the evaluation findings. Throughout this guide, we speak about evaluations as being conducted by independent third parties. That is distinct from monitoring activities which are typically conducted internally by Foundation program staff.

Planning

Planning is the most important and complex part of evaluation. Below are key steps and case examples that illustrate successes, pain points, and lessons learned.

Beginning evaluation design early

As part of the OFG process, a program team should consider the key assumptions in its theory of change and decide which warrant being systematically tested.

Often these are the assumptions that link the boxes in the causal chain of a logic model. For instance, consider this example of a simplified generic theory:

- If we invest in an innovative model, we hope and plan for it to be successful, and…

- if proven successful, it will be scaled to reach many more people.

In between each link are potential assumptions to be tested:

- This innovative approach can be successful.

- Effective organizations exist that can implement this approach.

- This approach can become a "model," and not just a one-off success.

Start evaluation planning early!

Six years after starting the ten-year **Special Initiative to Reduce the Need for Abortion,** Foundation staff began planning an evaluation whose primary purpose was to contribute to informing the staff and Board's future funding decision.

Designing an evaluation at this stage of implementation created challenges, some of which could have been minimized had an evaluation framework been established from the outset.

First, some of the long-term goals (e.g., reducing the number of abortions in the United States by 50 percent) do not now seem feasible and the "intermediate" targets are also high level and long term. If evaluative thinking had begun earlier, target setting might have been more realistic, and intermediate aims could have been identified and progress could have been measured in a systematic way.

Second, consultations with Foundation leadership during evaluation planning revealed an interest in answering questions about attribution (e.g., how much did this intervention cause the observed dramatic declines in the rate of teen pregnancy). However, the Initiative had not been designed to answer those questions.

Third, as a result, the evaluation was left to answer two questions at once, risking revisionist thinking: (1) what *would* have been possible for success at this point? and (2) how much progress has the Initiative actually made?

Key reflection: it would have been valuable to bring evaluative thinking to bear earlier in the process, as well as to allocate time and money for an evaluation from the start. The original evaluation plan would likely have needed modification over time, but still would have been a useful tool.

- Others will be interested in adopting and supporting the model.

- Resources for growth and expansion exist to scale the model.

As with many strategies, each link builds on the one before. So threshold evaluation questions that can help inform future direction are important to answer relatively early in the strategy's life. For instance, we might want to know first if an approach is effectively implemented and then if it is achieving desired outcomes before we advocate for scale.

This kind of evaluative thinking can help sharpen a theory of change from the outset, inform the sequencing of grantmaking, and highlight interdependencies to be supported or further explored.

Starting evaluation planning early in a strategy development process, rather than midway through an initiative, protects against four common pitfalls: (1) missing a "baseline"; (2) not having data available or collected in a useful common format; (3) surprised, unhappy, or unnecessarily burdened grantees; and (4) an initiative not optimally designed to generate the hoped-for knowledge.

Designing an evaluation framework does not mean casting in concrete. In fact, given that our strategies typically unfold dynamically, it is essential to revisit and modify an evaluation framework over time.

Clarifying an evaluation's purpose

The purpose of an evaluation is central. Questions, methods, and timing all flow from a clear understanding of how the findings will be used. Our three main purposes for evaluations are:

1. *To inform Foundation practices and decisions*. Evaluations with this aim may inform our decision making about funding or adapting an overall strategy, component, or initiative; setting new priorities; or setting new targets for results. These evaluations are typically designed to test our assumptions about approaches for achieving desired results.

2. *To inform grantees' practices and decisions*. At times, the Foundation may want to fund or commission evaluations of individual grantees or groups of grantees mainly to improve their practices and boost their performance. When the interests of the Foundation and grantees overlap, it may be worthwhile to commission evaluations of value to both. Collaborating in this way can promote more candor and buy-in for the ways data are collected and results are used. As necessary, we will support building our grantees' capacity to conduct evaluations and use the findings.

3. *To inform a field*. Sometimes evaluation itself can be part of a strategy—for example, to generate knowledge about what does and does not work in a field and why, and to have that knowledge shape its policy and practice. These evaluations, rigorously designed to achieve a high degree of certainty about the results, are usually shared widely.

The majority of our evaluations seek to inform the decisions and practices of the Hewlett Foundation and our grantees—to support our ongoing learning, adjustment, and improvement. The smaller number of evaluations we commission to inform broader fields are often intentional parts of program strategies and look more like research studies. Because they are often quite costly and long term in outlook, we commission these evaluations selectively and plan for them carefully.

For evaluations designed to inform Foundation decisions and approaches, it is important that we examine our level of openness to a range of results. Evaluation is worthwhile only if one can imagine being influenced by the findings. Are we willing to change strongly held beliefs in response to the evidence from an evaluation? If not, we should reconsider the value of spending money on it. If its purpose is to inform the Board and perhaps ongoing funding, are we clear on the Board's questions? Is the Board willing to change its strongly held beliefs?

For evaluations designed to inform grantees, we should consider how open and involved they are in the process. Do they have the capacity to devote to an evaluation? Are they driving it? If not, are they likely to abide by the results?

Evaluations intended to inform a field are usually fairly high stakes and meant to inform policy and significant resource allocation. Are we prepared for both positive and negative results (e.g., an intervention showing "no effect")? Are we prepared to share results with the field either way? Do we have a plan for influencing field decisions beyond passively posting an evaluation report?

Choosing what to evaluate

We cannot evaluate everything. Of course, a gating criterion for what we choose to evaluate is openness to change and readiness to challenge strongly held beliefs. Assuming that readiness threshold is met, several other criteria guide the decision about where to put our evaluation dollars. Highest priority is given to the following considerations:

- Opportunity for learning, especially for unproven approaches.

- Urgency for timely course correction or decisions about future funding.

- Risk to strategy, reputation, or execution.

- Size of grant portfolio (as a proxy for importance).

- Expectation of a positive expected return from the dollars invested in the evaluation.

Most of the time, especially when aiming to inform our decisions or a field's, an evaluation will focus on an initiative/component, subcomponent, or cluster of grants (grants that share some key characteristics, e.g., arts education grants) rather than on a single grant. The exception is when a grant is essentially

---

Challenging strongly held beliefs

In Mexico, the Environment Program conducted an evaluation of its Transportation portfolio in order to learn what had been accomplished, make a funding recommendation to the Board, and determine when to exit the different areas of work.

Surprisingly, one of the three strategies—the Clean Vehicles strategy—was shown to be more effective than the other two despite facing the strongest policy barriers. As a result, the team reallocated funding to this strategy and supplemented it with new policy angles and voices. At first, team members struggled to change their beliefs that the other strategies were not as effective (even in the face of fewer policy barriers), but they were convinced by the data and made decisions accordingly.

Choosing *not* to evaluate

In 2011, the Organizational Effectiveness (OE) Program decided against launching an evaluation of the Foundation's OE grantmaking. After careful consideration, the team determined that the costs of such an evaluation—including consultant fees, demands on OE grantees, and the significant OE and IT staff time needed to organize and analyze past grants data—would outweigh the anticipated benefit of the findings. At the same time, the Packard Foundation's OE Program, on which ours is largely based, was completing a comprehensive evaluation. Given the similarity between the two OE programs, our staff determined it was reasonable to draw conclusions about our grantmaking from the Packard Foundation's evaluation findings and leverage its lessons learned.

operating as an initiative or cluster in and of itself (e.g., The National Campaign to Prevent Teen and Unplanned Pregnancy or the International Development Research Centre's Think Tank Initiative).

It is most useful for a program to evaluate a whole strategy (initiative/component) at a reasonable mid-point and at its conclusion—to generate lessons that will be useful to multiple stakeholders inside, and potentially outside, the Foundation.

FOUNDATION'S CURRENT STRATEGY HIERARCHY

Program
|
Component or Initiative
|
Subcomponent
|
Grant cluster
|
Grant
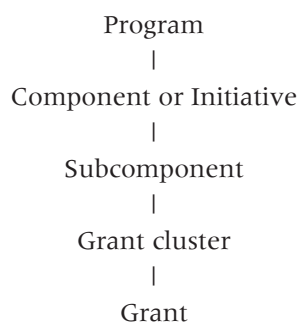
Frequently, the Foundation uses regranting intermediaries (e.g., through the ClimateWorks Foundation or the Community Leadership Project) to extend our reach and the impact of our grant dollars and results. Because we are delegating to these intermediaries what might be considered our stewardship role, we have an even greater responsibility to evaluate their efforts. By definition, large intermediaries rank high on the Risk and Size criteria above, and their evaluation typically offers important learning opportunities. Also, whenever we create a new intermediary organization or fund the launch of a major new initiative, it is important to evaluate not only the strategic elements but also issues of organization and effective execution, challenges that vex many start-ups. (For more on this subject, see the section on Special Evaluation Cases.)

Defining key questions

Our evaluations begin with and are guided by clear, crisp questions. Crafting a short list of precise questions increases the odds of receiving helpful answers—and a useful evaluation. Well-designed questions about an initiative or program can clarify not only the expected results but also surface assumptions about its design, causality, time frame for results, and data collection possibilities. These surfaced assumptions and questions can then help sharpen a theory of change and ensure effective planning for knowledge generation and learning.

Unfortunately, many evaluations begin to go awry when questions are drafted. It is useful to start by distinguishing between the following areas of inquiry. Although not every evaluation should seek to answer this full range of questions, the categories below offer a framework for effective investigation:

- **Implementation:** How well did we and our grantees execute on our respective responsibilities? What factors contributed to the quality of implementation?

  In much of the social sector, it is axiomatic that most programs fail in execution. This makes evaluating implementation very important for driving improvement, understanding the ingredients of a successful or failed approach, and replicating or adapting approaches over time.

- **Outcomes:** What changes have occurred? How do they compare with what we expected? To what extent and why are some people and places exhibiting more or less change? What is the relationship between implementation and outcomes?
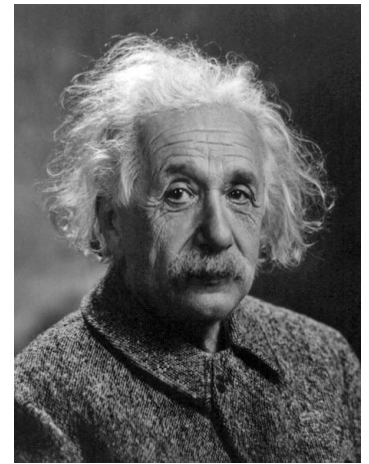

If I had an hour to solve a problem and my life depended on the solution, I would spend the first 55 minutes determining the proper question to ask, for once I know the proper question, I could solve the problem in less than five minutes.

–Albert Einstein

Limiting evaluation scope

As 300 million people moved from China's countryside to urban areas, the China Sustainable Cities Initiative aimed to shape this pattern of development so that Chinese cities focused on biking, walking, and mass transit rather than on car transportation.

The Environment Program decided to evaluate the Initiative because it was a significant investment and offered a great opportunity for learning. The original scope of the evaluation covered five areas:

1. Soundness of the theory of change.
2. Soundness of the Initiative's strategic plan.
3. Effectiveness of grantee implementation.
4. Possibility for replication.
5. Financial sustainability.

A key lesson for the Environment Program was that this first evaluation tried to cover too much ground. It proved most important to first answer basic questions about the Initiative's theory of change and execution, and then to use the results to make adjustments and shape future plans. Other questions about replication and sustainability were largely premature.

HYPOTHETICAL: "TEACHER AS LEARNER" INITIATIVE

Imagine that we are supporting a new initiative called "Teacher as Learner" that aims to improve the quality of teaching and learning in different regions in Africa via a network of 100 self-organized groups called "communities of practice." Each group of local teachers is professionally facilitated and focused on their specific capacity needs. Having organized themselves around issues of local importance, the "communities of practice" draw on regional resources as needed. The initiative's key assumption, based on some evidence in other fields, is that a blend of professional support and local ownership will lead to improved outcomes. If this approach seems successful after an initial period of innovation, we might develop an experiment to rigorously assess impact.

We did not plan for an evaluation at the beginning of the initiative because we were fatigued by the strategy development process and not quite sure how to focus an evaluation. Three years into this five-year initiative, we now want to commission an evaluation to see if we should adjust our approach.

Poor sample question: *Was the "Teacher as Learner" theory of change successful?*

This question has limited value for several reasons. First, it is vague. Usually a theory of change has multiple dimensions and contains many assumptions about how change will happen. A useful evaluation question is explicit about which interventions and assumptions it is exploring or interrogating. A vague question gives the evaluator too much discretion. This often sets us up for potential disappointment with the findings when we receive an evaluation report that is not useful and does not answer questions of importance to us. Or, it can set us up for an evaluation course correction—to shift midstream on design and data collection methods.

Second, and a related point: it is unclear whether the question is aimed at issues of execution (e.g., Did *x* happen?) or issues related to the "causal chain" of events (e.g., If *x* happened, did it catalyze *y?*). It is often useful in an evaluation to look at execution and outcomes with a distinct focus, as well as the relationship between them.

Third, the definition of success is unclear, allowing the evaluator too much discretion. Does success mean that 80 percent of what we hoped for happened? What if 60

percent happened? What if two out of three components progressed exactly as planned, but a third, delayed by an unforeseen personnel challenge, has not yet been implemented? Asking a dichotomous Yes/No question about an unspecified notion of "success" will be less helpful than a few focused questions that precisely probe what we want to learn and anticipate how we might use the answers.

Good sample questions

About implementation:

1. How and to what extent did the "Teacher as Learner" initiative create a network of local, self-organized communities of practice?

2. What was the nature of the variation in areas of focus for the communities of practice?

About intermediate outcomes:

3. To what extent did teachers adopt or adapt improved teaching methods after participating in the communities of practice?

4. What were the key factors that enabled or inhibited teachers from adopting new teaching methods?

About outcomes:

5. In what ways and by how much did these teachers' students improve their learning?

6. Is there any variation in students' learning gains? If so, what are possible explanations for that variation (including considerations of student, teacher, or community characteristics, and features and approaches used in the communities of practice)?

Why are these better questions? As a valuable beginning, they break one vague question about success into clear, specific ones that generate insight about different steps in the initiative's causal chain: which parts may be working well and as expected, which less well, and possible explanations why. They give more direction to the evaluator about our specific areas of interest. And, although they still need to be elaborated with specific measurement indicators and methods of data collection, they are designed to generate data that can be used to correct course.

To be able to answer these questions, it is enormously helpful to have planned an initiative's evaluation from the outset so that measurements are in place and changes are tracked over time.

- **Impact:** What are the long-term sustainable changes? To what can we attribute them?

  Although typically the most complex and costly to answer, questions that address long-term impact and attribution yield a significant return on investment when they can inform a field.

- **Context:** How is the landscape changing? Have changes in the world around us played an enabling or inhibiting role in our ability to affect change?

  Often our theories of change involve assumptions about how the world around us will behave, and unanticipated events—conflicts, new governments, social protests, disease, technological or scientific breakthroughs—can accelerate or slow progress toward our long-term goals. Understanding these interplays can help us avoid false conclusions.

- **Overall Strategy and Theory of Change:** Did our basic assumptions turn out to be true, and is change happening in the way we expected?

  Answering these questions will draw from other evaluations and go beyond them into the realm of what is now being referred to in the field as "strategy evaluation."

Timing: By when do we need to know?

One criticism of evaluation is that results often come too late to act upon. But that is in our control! *There are trade-offs to keep in mind, but it is important to* NOT sacrifice relevance by having evaluation findings be delivered too late to matter.

If we want to inform Foundation decisions, what is our timetable for seeing at least preliminary results? How firm is that timetable? Backing up from there, when would we need to have results in order to make sense of them and to bring them forward for funding considerations? If we want actionable information, it is essential to grapple with what is knowable in what time frame.

If we are aiming to inform grantees, how might their budgets or program planning cycles affect the evaluation timetable? Grantees also need time to make sense of findings and act upon them.

If our purpose is to inform the field, are there seminal meetings or conversations that we want an evaluation to influence? Are there election debates, planning processes, or budget cycles that might be important to consider in our evaluation planning? Many evaluations that target field-level change benefit from some legal consultation to ensure appropriate engagement.

Of course, considering evaluation early as part of strategy development will help define when specific information will be needed.

Selecting methods

Most strong evaluations use multiple methods to collect and analyze data. This process of triangulation allows one method to complement the weaknesses of another. For example, randomized experiments can determine whether a certain outcome can be attributed to an intervention. But complementary qualitative methods are also needed to answer questions about how and why an intervention did or didn't work—questions that are central to replication. Thus, as part of early planning, it is ideal to select methods that match evaluation questions.

Our goal is to maximize rigor without compromising relevance. Part of maximizing rigor is reducing bias in the evaluation. While not all evaluations can feasibly be randomized so that we can definitely attribute impact to one or more interventions, the essence of good evaluation involves some comparison—against expectations, over time, and across types of interventions, organizations, populations, or regions. Even when there is no formal counterfactual, it can be helpful to engage in "thought experiments" to challenge easy interpretations of data and consider alternative explanations.

{ *The essence of good evaluation involves some comparison—against expectations, over time, and across types of interventions, organizations, populations, or regions.* }

Multiple methods help reduce bias as does active consideration of how the methods are applied. For instance, if an advocacy initiative is being evaluated largely through qualitative interviews of key informants, it will be important to include respondents who are not cheerleaders, but may offer constructive critiques.

Engaging with grantees

It is essential that Foundation staff engage with grantees about evaluation and communicate with them early and often about expectations. What is communicated and how will of course depend on the purpose of the evaluation and the grantee's role in it. At a grant's inception, program staff should inform grantees that they may be expected to participate in an evaluation, share data with the Foundation and evaluators, and potentially, if relevant, have the results shared with the field (see Appendix B). It is never a good idea to surprise grantees with an evaluation. Often this expectation needs to be communicated and reinforced several times. As one grantee who reviewed this guide advised us, "Don't sugarcoat what the evaluation experience will entail." In the long run, everyone does better when expectations are clear.

Another reviewer said, "The relationship between the evaluators and the implementers is KEY" to successfully conducting an evaluation and applying the findings. If grantees are engaged about the evaluations that touch them, they will be: (1) more supportive with respect to data collection; (2) more likely to learn something that will improve their work; (3) less likely to dismiss the evaluation; and (4) better able to help strengthen the evaluation design, especially if engaged early. From a design perspective, this last point is quite

important: grantees can serve as a reality check and deepen understanding of the available data and data collection systems.

Crafting an RFP for an evaluator

The basic elements of an RFP to engage an evaluator include background information about the evaluation, its purpose, key evaluation questions, known available data sources, time frame for receiving results, intended audiences, preferred deadline for the deliverable, and amount of available funding. For examples of different ways to craft these documents, click here.[3]

Choosing an evaluator and developing an agreement

The ideal evaluator is strong technically, has subject matter expertise, is pragmatic, *and* communicates well, both verbally and in writing. Often in our work, cultural

---

Choosing an evaluator

The Communications Department conducted an evaluation of its Communications Academy, an intensive, multiday training for leaders from our grantee organizations. Under both time and budgetary restraints, staff had to choose between a firm with subject matter expertise and one with evaluation expertise, since they could not find an evaluator who possessed both. They ultimately chose communications experts. Although the team was pleased with the evaluation and used its findings to refine their program, it was less rigorous than their ideal. In retrospect, the Communications staff would have liked to create a "dream team" of individuals with both types of expertise.

The Education Program is conducting a three-year evaluation of its **Deeper Learning Network**. This initiative aims to create a coalition of K-12 schools to act as beacons of effective practice for building students' abilities to think critically, solve complex problems, and learn how to learn. After a rigorous planning process, the Program invited three leading research firms to submit evaluation proposals. To ensure broad support and buy-in from Network leaders, program staff solicited their detailed comments about the research design and methodology, along with feedback from research experts. These comments informed the team's choice of the best consultant to manage and implement this evaluation.

Building trust with grantees

The Community Leadership Project was launched in April 2009 as a $10 million funding partnership between the Packard, Irvine, and Hewlett foundations. Its purpose is to build the capacity of small and midsize nonprofits serving low-income people and communities of color in three California regions. The three foundations made grants to twenty-seven intermediary organizations that in turn regranted funds to community organizations and provided their leaders with technical assistance to strengthen a wide range of their skills and abilities.

The funders were interested in an evaluation that would assess the effectiveness of the overall project and also prove useful to the grantees.

Evaluation challenges. Because the evaluation was designed after the project was launched, it surprised the grantees. They were initially very resistant: participating would be costly in terms of their time (many organizations had just one or two staff members) and labor (participants would have to travel long distances to convene). In addition, some assessments seemed to duplicate internal evaluations the nonprofits had already developed.

Also, because the proposed organizational capacity building included support for grantees' cultural competency, it was important that the project evaluation likewise demonstrate sensitivity to cultural diversity as a key principle and concern.

Course corrections. Based on grantees' feedback, the funders made a number of course corrections during the evaluation. They simplified its methodology, coordinated with existing assessments, and reduced the number of required meetings and frequency of required reporting. They provided reimbursements for participants' time and travel expenses. And they hired an evaluator experienced in and sensitive to issues of culture in the communities.

Lessons learned. By responding to grantees' feedback, the funders encouraged their active engagement in the evaluation. Real-time learning allowed continual course corrections and proved more valuable than a summative evaluation. The evaluation consultant's cultural competency, an even more important asset than the funders originally anticipated, went a long way toward building trust with grantees.

---

[3] This link is for internal use only. People outside of the Foundation will not be able to access it.

awareness and sensitivity to the context in which nonprofits are operating are also very important. If we cannot find that full package, it may sometimes be appropriate to broker such a marriage and bring people or teams together with complementary skills. Choices always involve trade-offs; it is important to manage their risks.

Engaging grantees can be helpful in the evaluator selection process. They not only become invested in the effort but also often contribute a useful pragmatic perspective.

When developing a contract or grant agreement with an evaluator, be sure to address questions of data ownership. Will we own the data and the research findings? Will the evaluator own them? Or will the grantee own them?

Implementation

More often than not, an evaluation's implementation does not go precisely as planned. Staying connected with the evaluator and the evaluation during implementation can go a long way towards ensuring responsiveness and a generally higher quality evaluation.

Managing the evaluation

Active management is essential. Talk with the evaluator regularly and ask what is or is not going well. Request periodic memos to document progress and any obstacles the evaluator is facing in data collection, data quality, or other areas. These exchanges can be useful forcing functions to keep an evaluation on track and to start troubleshooting early. Often the data collection in an evaluation mirrors some of the challenges faced by a program in other facets of its work, so evaluation progress updates can be helpful in multiple ways.

It can be especially useful to set an expectation of interim evaluation reports on preliminary findings or baseline data summaries. This will keep an evaluation on course, engage Foundation staff in the loop of learning, and allow any needed course corrections.

Responding to challenges

Not surprisingly, the best laid plans of an evaluation do not always unfold as designed on paper. Any number of challenges can emerge: a data source may be less reliable than predicted; survey response rates may be too low to draw conclusions; other interventions may have developed that make a planned comparison group suboptimal; or staff turnover in the selected firm may reduce confidence in the actual evaluation team.

If you hit these bumps or others in the evaluation road, it is important to pause, take stock of the challenges, revisit prior plans, consult appropriate stakeholders, consider alternative solutions, and make necessary course corrections.

Bumps in the evaluation road

A program officer recently commissioned an evaluation of her initiative to inform a funding recommendation to the Board. She hired an expert in the field as the evaluator, considering him best able to understand the field's complexities, nuances, and players.

Midway through the evaluation, she realized that this expert was strongly biased and could not objectively assess a community of practitioners to which he belonged. Although the evaluation was well under way, the program officer moved quickly to rectify the problem. She engaged an independent evaluator to work with the field expert, successfully gaining both the objectivity and the expertise she needed for a rigorous, relevant evaluation.

Don't forget to close the loop and communicate any changes to everyone invested in the work, including grantees.

Synthesizing results at the strategy level

Most strategies operate with several initiatives, often with multiple clusters nested in each. Typically, no single evaluation can tell us if a strategy has been successful or is on track. Such a comprehensive assessment requires synthesis of multiple evaluations, summary and analysis of relevant performance indicators, and active reflection on and interpretation of the results in context. This process can be more of a quilting art than an exact science. There is value in having a third party assist with such an evaluation to increase objectivity. However, strategy evaluation is a relatively new area in the evaluation field, and there is no consensus about what it should look like.

Using Results

Using results is often messier than anticipated. Sometimes, staff expect more confirmation of success than an evaluation typically delivers. Sometimes, an evaluation is not especially well done, and the results inspire limited confidence. Sometimes, staff simply do not know how to apply the lessons. They are uncertain how best to shift a strategy or overhaul a program.

From the very beginning of the evaluation process, it helps tremendously to plan how the results will be used; along the way, it is wise to remind yourself of those intended uses. Staffing changes—whether within the Foundation, the evaluation team, or a key grantee organization—present important moments to recommit to an evaluation and its intended uses. Often a short exercise of predicting the findings can helpfully surface assumptions about them and generate discussion about what might be done differently if these assumptions are not borne out.

Taking time for reflection

If in the beginning, you take time to imagine how you might respond to different results scenarios, you are halfway toward actual use of the findings!

Take time for reflection and the development of insights. Ask questions of the evaluators, grantees, your colleagues, yourself. Make sure you don't get all the way across the finish line of an evaluation and just file the report on a shelf or in a drawer. It is amazing how often this happens—partly because many evaluations have not been adequately planned with purpose, relevant questions, audiences, and timing considerations in mind.

Sharing results internally

Sharing the results of an evaluation with Foundation colleagues brings many benefits, and it is worthwhile to build this step into your process. For staff managing an evaluation, these discussions can crystallize the results, lead to

Ways to think about using the data

▶ Convene grantees to discuss the results and recommendations

▶ Organize an internal briefing to share with your colleagues what you've learned, both about your strategy projects and the evaluation process itself

▶ Discuss with your board how the evaluation results will inform changes in your strategy or grantmaking approach

▶ Share a version of the evaluation (e.g., executive summary) with the field, accompanied by a memo detailing how you are applying the findings in practice

a deeper grasp of them, and force some grappling with what is not yet understood. It can also help staff think through what the results mean programmatically and how to apply them in practice. For members of other teams, review of the results can generate insights about their own programs, grantmaking approaches, or evaluation designs. An internal debrief at the conclusion of each evaluation about what went well and what did not, key lessons learned, and actions taken will help advance evaluation practice at the Foundation and keep us focused on designing evaluations with action in mind.

If another funder has collaboratively supported the evaluation, it is often appropriate to consider that partner an internal colleague with respect to sharing results and surfacing implications.

Sharing results externally

Our intention is to share evaluation results—both the successes and failures—so that others may learn from them. Out of respect, we communicate with our grantees early on about our intention to evaluate and listen to any concerns they may have about confidentiality. Grant agreement letters specify the organization's participation in an evaluation, clarify its purpose (including any anticipated effect on the grantee), the process for making decisions about it, and the roles for each party's participation. We also strike an agreement regarding the level of findings (full evaluation results, redacted evaluation results [no grantee names], or executive summary) that will be shared with which audience.

On principle, we consider the question of sharing evaluation findings on a case-by-case basis, with care given to issues of organizational confidentiality. For instance, if an evaluation is in part focused on questions of organizational development, it may be more useful for the findings to be shared only with that grantee, so it may use the results to drive improvement without having to take a defensive public stance.

Appendix C offers an internal planning tool for sharing results. It is designed to help program staff think about this process early on and consider implications for grantee organizations, requested reports, budgets, and communication plans.

The Foundation supports a considerable amount of grantmaking that involves regranting intermediaries, advocacy, and organizational capacity building. Over time, this Practice Guide will be expanded to cover evaluation of each of these grantmaking areas. We will develop special sections on these and other issues that have emerged in our own practice and which we discuss and explore in a series of special program staff meetings.

Evaluating Regranting Intermediaries

Every program at the Hewlett Foundation uses regranting intermediaries to help implement their grantmaking strategies and achieve their goals—for example, the Performing Arts Program, whose intermediaries support individual artists; the Environment Program, which funds the Energy Foundation to manage grant portfolios that promote clean energy; and the Education Program, which engaged Rockefeller Philanthropic Advisors in 2010 to invest in organizations whose work advances state policies that better support student achievement.

{ *A **regranting intermediary** is a charity that regrants funds under certain programmatic guidelines and often provides additional services such as capacity-building.* }

Evaluating the Foundation's regranting intermediaries is worth highlighting because of their prevalent use, the high dollar amount usually involved, and the complexities of this type of relationship. In addition, because these intermediaries carry forward our program strategies and reduce the programmatic dollars we directly provide grantees, it is important to assess their effectiveness and identify areas for improvement.

Evaluating a regranting intermediary requires a key element beyond all those already outlined in this guide: measuring the added value of the intermediary itself. To do this, you might seek to answer questions like these: How and to what extent is the intermediary adding value to its grantees? Is it just a middleman, supporting the transaction of regranting funds without adding much additional value? Or, is it able to offer important technical assistance to organizations by virtue of being closer to the ground? Where and with whom is the intermediary adding the most value, and where is it adding the least? What are the enablers and inhibitors to an intermediary's high performance? How does this intermediary's performance compare to others?

The following case study about the Think Tank Initiative is a worthy example because of its complexity: multiple funders, multiple levels of evaluation inquiry, and an internationally focused strategy and implementation plan.

Think Tank Initiative

The Think Tank Initiative is a ten-year, $100 million effort launched in 2009 with support from five major funders, including the Hewlett Foundation's Global Development and Population Program (GDP). The Initiative aims to strengthen independent research centers in the developing world so that their high-quality work can be used to formulate sound national policies. One funder also serves as the regrantor and implementation manager: the International Development Research Centre (IDRC). This quasi-public Canadian agency has forty years of experience supporting research on development.

The funders jointly agreed to commission an independent, external evaluation of the Initiative's first five years in order to inform decisions about the funding and design of the anticipated five-year Phase Two.

In planning for this evaluation, GDP staff quickly realized that it would be a complex undertaking for several reasons.

Establishing the evaluation framework. As planning began, it became apparent that the funders had different perspectives on which evaluation questions should take priority, initially resulting in a rather expansive evaluation framework. The evaluation team then struggled to design an approach to address the multitude of questions within the given timeframe and budget. Rather than propose ways to focus the framework, the team tried to incorporate everyone's input and concerns, which led to a muddled evaluation plan with far too many objectives.

Clarifying the different levels of evaluation. The funders ultimately concluded there was a need to address three different levels of inquiry in the assessment (illustrated by these simplified questions):

- **Initiative design:** How and to what extent did our theory of change play out in practice?

- **Program implementation:** How successfully did IDRC provide the needed support to its grantees to ensure their success?

- **Think tank impact:** In what ways and to what extent did the funded think tanks improve their performance and influence over the course of the grant period?

Still now, some funders are laser-focused on assessing the impact of individual think tanks, while others are satisfied with synthesized findings and a sampling approach that addresses the Initiative's overall success. However, this framework has proved to be a critical anchoring point for the evaluation. It enables recognition of three areas of inquiry and analysis and the different methodologies required for each.

Structuring and managing the process. Initially, because of IDRC's extensive institutional expertise with evaluation, the funders asked it to manage the selection of a team of experts to run the external evaluation. This turned out to be problematic for two reasons. First, having the implementing organization in charge of the evaluation did not ensure that it would be truly conducted at arm's length. Second, given the evaluation's scale and complexities, it was ineffective to piece together a team of individuals who had not worked together in the past, were not physically colocated, and did not have the required program management skills and experience.

As a result, the entire approach was revised. The funders' Executive Committee decided it should directly commission and manage the evaluation, nominating a committee member as the point person to coordinate the evaluation team selection process. The funders worked together to revise the evaluation's scope, questions, and clarity of deliverables. They selected an independent, collaborative evaluation team from two institutions that have partnered on a number of projects in the past, each bringing unique core competencies. And, they explicitly articulated and agreed on an evaluation plan that includes the timely submission of deliverables and ongoing communication between the funders and the implementer (IDRC).

The evaluation is currently proceeding on track, having significantly reduced the barriers to success. The Think Tank Initiative team is relatively confident that the final evaluation reports will be rigorous, nuanced, reliable resources that can guide decisions about the Initiative's future direction.

APPENDIX A   GLOSSARY

Activities. The actions taken by the Foundation or a grantee to achieve intermediate outcomes and make progress toward the achievement of goals. [Gates Foundation glossary]

Baseline. An analysis or description of the situation prior to an intervention, against which progress can be assessed or comparisons made. [Gates Foundation glossary]

Cluster. A small group of grants with complementary activities and objectives that collectively advance a strategy toward its goal.

Component. The different areas of work in which a program decides to invest its resources in order to achieve its goals. Each component typically has its own theory of change, logic model, strategy, and progress indicators and outcomes—all of which are designed to advance the program's overall goals.

Evaluation. An independent, systematic investigation into how, why, and to what extent objectives or goals are achieved. It can help the Foundation answer key questions about grants, clusters of grants, components, initiatives, or strategy. [Variant of Gates Foundation glossary]

> **Impact Evaluation**. A type of evaluation design that assesses the changes that can be attributed to a particular intervention. It is based on models of cause and effect and requires a credible counterfactual (sometimes referred to as a control group or comparison group) to control for factors other than the intervention that might account for the observed change. [Gates Foundation glossary; USAID Evaluation Policy]
>
> **Performance Evaluation**. A type of evaluation design that focuses on descriptive or normative questions. It often incorporates before/after comparisons and generally lacks a rigorously defined counterfactual. [USAID Evaluation Policy]
>
> **Formative Evaluation**. An evaluation that occurs during a grant, initiative, or strategy to assess how things are working while plans are still being developed and implementation is ongoing. [Gates Foundation glossary]
>
> **Summative Evaluation**. An evaluation that occurs after a grant or intervention is complete in order to fully assess overall achievements and shortcomings. [Gates Foundation glossary]
>
> **Developmental Evaluation**. A "learn-by-doing" evaluative process that has the purpose of helping develop an innovation, intervention, or program. The evaluator typically becomes part of the design team, fully participating in decisions and facilitating discussion through the use of evaluative questions and data. [Variant of *The Encyclopedia of Evaluation* (Mathison, 2005) and *Developmental Evaluation* (Quinn Patton, 2011)]

Evidence. A general term that refers to qualitative and quantitative data that can inform a decision.

Goal. A clearly defined, specific, achievable, and measurable program outcome that is broad enough to capture long-term aspirations but tailored to be achievable with the program's resources. It includes both a rationale and scope.

> **Rationale**. An explanation of why a specific goal is important and what distinct advantages and capabilities the Foundation could bring to bear in addressing it.

> **Scope**. A description of the geographies, topics, or other targets where grantmaking will be focused in order to make the greatest marginal impact with program resources, or to align with Foundation values. [OFG Overview]

Grant. A sum of money used to fund a specific project, program, or organization, as specified by the terms of the grant award.

Indicators. Quantitative or qualitative variables that specify results for a particular strategy, component, initiative, subcomponent, cluster, or grantee. [Gates Foundation glossary]

Initiative. A time-bound area of work at the Foundation with a discrete strategy and goals. Initiatives reside within a program, despite occasionally having goals distinct from it (e.g., the Think Tank Initiative within the Global Development and Population Program and the Nuclear Security Initiative within the Environment Program).

Inputs. The resources used to implement activities. [Gates Foundation glossary]

Logic Model. A visual graphic that shows the sequence of activities and outcomes that lead to goal achievement. [OFG Overview]

Metrics. Measurements that help track progress throughout a grant, component, initiative, or strategy.

Monitoring. A process of tracking the milestones and progress of a grant or portfolio against expectations, for purposes of compliance and adjustment. Evaluation will often draw on grant monitoring data but will typically include other methods and data sources to answer more strategic questions.

M&E. An acronym used as shorthand to broadly denote monitoring and evaluation activities. It includes both the ongoing use of data for accountability and learning throughout the life of a grant, component, initiative, or strategy, as well as an examination of whether outcomes and impacts have been achieved. [Gates Foundation glossary]

Outcome-focused Grantmaking (OFG). A structured, strategic approach to grantmaking designed to achieve the most impact with philanthropic funds. It is based on strategic planning principles from both business and philanthropy. [OFG Overview]

Outcomes. Results or change, ideally observable and measurable, based on a set of inputs and activities. Outcomes can be intermediate (e.g., results that serve as steps toward a goal) or ultimate (e.g., the goal).

Strategy. A plan of action designed to achieve a particular goal.

Targets. The desired level for goals the program plans to achieve with its funding. They are based on metrics and should be ambitious but achievable within the specified time frame.

Theory of Change. A set of assumptions that describe the known and hypothesized social and natural science underlying the graphic depiction in a logic model. [OFG Overview]

APPENDIX B   EVALUATION CONSENT IN GRANT AGREEMENT LETTERS

The Hewlett Foundation tries to be as transparent as possible to its grantees about its desire to conduct evaluations, either of the grantee itself or of a larger initiative to which the grantee's work contributes. To this end, Foundation grant agreement letters (essentially, our legal contract with an organization that receives funding) now include language stating our expectation that grantees will cooperate with Foundation evaluation efforts, as requested.

Default language for all grant agreement letters:
The Foundation may choose to conduct an evaluation of the effectiveness of this grant (the "Evaluation") either individually or as part of a broader Foundation strategy. Grantee agrees to cooperate in the Evaluation and provide such information to the Foundation or its representatives as necessary.

Grantee further agrees that the Foundation can disseminate to the public the results of the Evaluation, including any data created in connection with the Evaluation. In such cases, the Foundation may share the results of the Evaluation with the Grantee and may provide an opportunity for the Grantee to comment.

At the request of program staff, this default language may be changed to one of these alternatives:

1. **(Executive Summary)** Grantee agrees that the Foundation may disseminate to the public an executive summary of the results of the Evaluation. In such cases, the Foundation may share such executive summary with the Grantee and may provide an opportunity for the Grantee to comment.

2. **(Redacted Version)** Grantee agrees that the Foundation may disseminate to the public a redacted version of the Evaluation, including summary data created in connection with the Evaluation. In such cases, the Foundation may share the redacted version with the Grantee and may provide an opportunity for the Grantee to comment.

3. **(Internal Use Only)** The Foundation may share the results of the Evaluation internally with staff and consultants. The Foundation will not share the results of the evaluation publicly without the written permission of the Grantee.

APPENDIX C   PLANNING TOOL: SHARING RESULTS

Consistent with our newly adopted *"Evaluation Principles and Practice"* guide, we will consider and identify audiences for our evaluation findings early in the process, during the planning stage. We will presumptively share the results of our evaluations so that others may learn from our successes and failures. We will make principled exceptions on a case by case basis, with care given to issues of confidentiality and supporting an organization's improvement.

This internal tool is intended to help staff be intentional about the audiences with which we plan to share specific evaluation results, and articulate the rationale for cases where we plan to limit the distribution of evaluation results.

Grantee/Cluster/ Initiative/Strategy:				
Evaluation Purpose:				
WHO? (Audience)	**WHAT? (Product*)**	**WHY? (Communications strategy/purpose)**	**HOW? (Messenger)**	**NOTES (e.g., lead PO, timing, sequencing)**
Default (transparency)				
General public		-	HF website	
HF program colleagues		-		
HF administrative colleagues		-		
Optional (strategic outreach)				
HF Board				
Grantees evaluated				
Evaluation interviewees				
Co-funders				
Other funders in the space				
Funder affinity groups				
Grantee affinity groups				
Field opinion leaders, government officials, academia				
Media				

* The "product" could consist of: (1) Full evaluation results; (2) Redacted version (no identifying organizational names or characteristics, or other confidential information); or (3) Executive summary only (a high level summary of findings).

APPENDIX D   ACKNOWLEDGMENTS